

Reinhard Fiehler

Thesen zur Struktur einer Grammatik der gesprochenen Sprache

1 Einleitung

Brauchen wir eine Grammatik der gesprochenen Sprache? Es liegt auf der Hand, dass die Antwort auf diese Frage unterschiedlich ausfällt, je nachdem was man auf der einen Seite unter ‚Sprache‘ und auf der anderen Seite unter ‚Grammatik‘ versteht. Konzeptualisiert man Sprache als ein einheitliches Sprachsystem, das sowohl für die schriftliche wie für die mündliche Verständigung grundlegend ist, wird man eher dazu neigen, diese Frage zu verneinen. Sieht man schriftliche und mündliche Verständigung als unterschiedliche Verständigungsformen, die je spezifischen Bedingungen unterliegen und dementsprechende Regularitäten ausgebildet haben, wird man eher dazu tendieren anzunehmen, dass diese Besonderheiten ein Ausmaß besitzen, das es wert ist, auch grammatisch beschrieben zu werden. Aber damit nicht genug: Wenn man denn der Meinung ist, dass es eine solche Grammatik geben sollte, stellen sich sofort die Anschlussfragen, wie sie aussehen soll und was sie zu beschreiben habe, ob sie nur die gesprochene Sprache darstellen oder ob es eine integrierte Grammatik sein soll, die gesprochene und geschriebene Sprache im Zusammenhang behandelt. Über diese Fragen ist theoretisch viel gestritten worden und ließe sich auch weiterhin streiten, gäbe es nicht von Zeit zu Zeit handlungsorientierende Impulse aus der Praxis.

Die Dudenredaktion auf jeden Fall kam, als sie im Jahr 2000 die 7. Auflage der Duden-Grammatik plante, zu der Auffassung, dass diese Neuauflage um Kapitel zur Textgrammatik und zur Grammatik gesprochener Sprache im Umfang von jeweils achtzig Druckseiten erweitert werden sollte:

Als dritte wichtige Änderung schlagen wir vor, einen eigenen Abschnitt zur Grammatik der gesprochenen Sprache aufzunehmen. Dies berücksichtigt nicht nur den Forschungsfortschritt in der Germanistik, sondern kann zugleich Normunsicherheiten von Laien klären helfen, die sich aus der Diskrepanz zwischen geschriebener und gesprochener Sprache ergeben. Wir meinen, dass die Dudenredaktion im Bereich der Grammatik auch über ihren eigentlichen Gegenstand (die *geschriebene* deutsche Standardsprache der Gegenwart) hinausgehen sollte, damit die Duden-

grammatik in Zukunft eine wirklich moderne Grammatik sein kann. (DUDEN 4 – Die Grammatik, Konzept für die 7. Auflage, S. 13)

Als ich Anfang 2001 die Aufgabe übernahm, das Kapitel zur Grammatik gesprochener Sprache zu schreiben, waren damit deutliche Rahmenbedingungen gesetzt und die Grundsatzfrage pragmatisch entschieden. Umso drängender stellten sich aber Fragen wie: Was muss man, was kann man in welchen Proportionen auf achtzig Seiten über gesprochene Sprache darstellen? Wie kann es in den Rahmen der dominant schriftsprachlich orientierten Grammatik eingefügt werden?

Der sich anschließende Klärungsprozess hat mich zu einer Reihe von Annahmen und Entscheidungen geführt, auf deren Basis ich dann eine Konzeption und Gliederung für das Kapitel entwickelt habe. Um diese Hintergründe zu verdeutlichen, habe ich meine Annahmen und Entscheidungen im folgenden Abschnitt in Form von sieben Thesen zusammengefasst, die ich dann jeweils im Anschluss eingehender kommentiere.

2 Die Thesen

These 1: Aufgaben der Grammatik

Gegenstand der Grammatik sind die *Einheiten* einer Sprache. Sie bearbeitet dabei drei Aufgaben:

- (1) Sie beschreibt die *grundlegenden Einheiten* einer Sprache und deren Eigenschaften.
- (2) Sie beschreibt die Regularitäten des *Aufbaus* dieser Einheiten.
- (3) Sie beschreibt die Möglichkeiten der *Verknüpfung* von Einheiten.

Sie hat die dabei jeweils geltenden *Regularitäten* in *systematischer Weise* darzustellen.

Anmerkungen zu These 1:

Es gibt vielfältige andere Auffassungen darüber, was Aufgabe einer Grammatik sei:

- Grammatik stellt fest, was zu einer Sprache gehört und was nicht.
- Grammatik beschreibt die Regeln einer Sprache.
- Grammatik beschreibt das Sprachsystem (*langue*).
- etc.

Dies ist nicht der Ort, diese vielfältigen Auffassungen zusammenzustellen und zu diskutieren, inwieweit und worin sie sich unterscheiden. Es soll nur festgehalten werden, dass unterschiedliche Vorstellungen darüber, was Aufgabe der Grammatik ist, zu unterschiedlichen Ansichten führen, wie eine Grammatik der gesprochenen Sprache zu konzipieren ist – bis hin zu der Auffassung, dass man eine solche nicht brauche oder dass es sie nicht geben könne.

Die folgenden Überlegungen basieren auf der in These 1 gegebenen Aufgabenbestimmung von Grammatik.

These 2: Einheiten geschriebener und gesprochener Sprache

Grundlegende Einheiten der geschriebenen Sprache sind – nach zunehmender Größe geordnet:

- Der *Buchstabe*. Er ist in Alphabetschriften die elementare Konstruktionseinheit.
- Das *Wort*. Es ist das zentrale Element, um sprachlich auf die Welt Bezug zu nehmen.
- Der *Satz*. Er ist die elementare Aussageeinheit.
- Der *Text*. Er ist die in sich abgeschlossene, schriftliche kommunikative Einheit.

Fragt man nach den entsprechenden grundlegenden Einheiten der gesprochenen Sprache, so sind die folgenden zu nennen:

- Der *Laut*.
- Das *Wort*.
- Die *funktionale Einheit*.
- Der *Gesprächsbeitrag*.
- Das *Gespräch*.

Anmerkungen zu These 2:

Während über die grundlegenden Einheiten der geschriebenen Sprache weitgehend Konsens bestehen dürfte, fällt die Antwort auf diese Frage bei der gesprochenen Sprache weit weniger einheitlich aus. Entsprechend sind sehr unterschiedliche Grundeinheiten benannt worden (s. Anmerkungen zu These 7).

Versucht man, die gesprochene Sprache nach dem Modell der schriftlichen zu verstehen, so kann man die genannten schriftsprachlichen Einheiten übertragen bzw. Analogisierungen vornehmen. Den Buchstaben entsprechen

dann die Laute, dem geschriebenen das gesprochene Wort und dem Text als kommunikativer Einheit das Gespräch. Diese Parallelisierung erscheint zunächst problemlos – bis auf den Punkt, dass das Gespräch nicht in gleicher Weise aus Sätzen besteht, wie der Text es tut. An diesem Punkt entzieht sich die gesprochene Sprache aufgrund ihres Eigencharakters einer einfachen Analogisierung. An die Stelle des Satzes treten der Gesprächsbeitrag, den Sprecher abwechselnd beim Führen eines Gesprächs leisten, und funktionale Einheiten, aus denen sich die einzelnen Gesprächsbeiträge zusammensetzen (s. These 7).

Die ‚größeren‘ Grundeinheiten der gesprochenen Sprache – funktionale Einheit, Gesprächsbeitrag und Gespräch – sind dabei nicht nur strukturelle Einheiten, sondern haben zugleich auch einen stärker handlungsbezogenen Aspekt als die Grundeinheiten der geschriebenen Sprache.

These 3: Aufgaben einer Grammatik gesprochener Sprache

Eine Grammatik gesprochener Sprache hat die fünf genannten grundlegenden Einheiten zu beschreiben, die Regularitäten ihres Aufbaus darzustellen und die Möglichkeiten ihrer Verknüpfung auszuarbeiten.

Anmerkungen zu These 3:

Eine Grammatik gesprochener Sprache hat damit zu einem großen Teil *andere* Grundeinheiten zu beschreiben als die der geschriebenen Sprache: Dies sind der Laut, die funktionale Einheit, der Gesprächsbeitrag und das Gespräch. Bei der gemeinsamen Einheit ‚Wort‘ ist zu prüfen, inwieweit in der gesprochenen Sprache aufgrund ihrer lautlichen Struktur andere Regularitäten bestehen (z. B. lautliche Veränderungen im Vergleich mit der an der Schreibung orientierten Explizitlautung) und inwieweit es eine spezifische Verteilung von Wörtern bzw. Wortklassen gibt (z. B. ‚Gesprächswörter‘, Abtönungspartikeln). Bei der ‚funktionalen Einheit‘ ist herauszuarbeiten, welche nicht satzförmigen Einheiten für die gesprochene Sprache wesentlich sind, und bei den satzförmigen funktionalen Einheiten ist festzustellen, welche Unterschiede zu den Sätzen der geschriebenen Sprache bestehen.

These 4: Integrative Grammatik

Die Grammatik einer Sprache sollte eine integrative Grammatik sein, die intern aber eine deutliche und systematische Aufteilung in eine Grammatik der geschriebenen Sprache und eine Grammatik der gesprochenen Sprache aufweist.

Anmerkungen zu These 4:

In Gesellschaften mit einer entwickelten Schriftsprache sind gesprochene und geschriebene Sprache in vielfältiger Weise miteinander verwoben und eine grammatische Beschreibung der Sprache sollte beide Bereiche in angemessener Weise berücksichtigen.

Zu beachten ist dabei der unterschiedliche Entwicklungsstand der Grammatikschreibung in den beiden Bereichen: Die Erforschung der gesprochenen Sprache hat – verglichen mit der an der Schriftlichkeit orientierten Grammatikschreibung – eine vergleichsweise kurze Tradition, die kaum älter als 100 Jahre ist. Behaghel (1899) kann man dabei als Startpunkt sehen. Entsprechend hat die Beschreibung der gesprochenen Sprache und ihrer Grammatik noch keine kanonischen Standards entwickelt, sondern die Ausarbeitung von Beschreibungskonzepten und -kategorien ist in einer ständigen Entwicklung begriffen.

Die Beschreibung gemeinsamer Regularitäten, die sowohl für die geschriebene als auch für die gesprochene Sprache gelten, verbleibt in einer integrativen Grammatik aus historischen Gründen eher bei der Grammatik der geschriebenen Sprache, so dass die Grammatik der gesprochenen Sprache bei der gemeinsamen Einheit ‚Wort‘ und partiell auch bei den funktionalen Einheiten den Charakter einer *Differenzgrammatik* besitzt, bei den genuinen Einheiten der gesprochenen Sprache hingegen aber *eigenständig* ist.

These 5: Die faktische Leistung von Grammatiken

Wenn man davon ausgeht, dass es – wie dargestellt – Aufgabe der Grammatik ist, die grundlegenden Einheiten einer Sprache und ihre Regularitäten in systematischer Weise zu beschreiben, so erfüllen viele der vorliegenden Grammatiken diese Aufgabe nur teilweise – sowohl für das Schriftliche wie auch für das Mündliche. Traditionelle Grammatiken sind im Wesentlichen Grammatiken der Schriftsprache, wobei auch hier in der Regel nicht alle Einheiten behandelt werden.

Anmerkungen zu These 5:

Im Mittelpunkt der meisten Grammatiken stehen das Wort und der Satz. Einige Grammatiken behandeln – in der Regel deutlich kürzer – die Buchstaben und den Text.

Von den Einheiten der gesprochenen Sprache werden lediglich die Laute (meistens im Zusammenhang mit den Buchstaben) ausführlich behandelt.

Bei den Einheiten ‚Wort‘ und ‚Satz‘ werden Phänomene der gesprochenen Sprache nicht eigenständig, sondern nur punktuell als Ergänzung oder Abweichung von den Verhältnissen in der geschriebenen Sprache thematisiert, so z. B. lautliche Besonderheiten der gesprochenen Sprache (Wegfall von Lauten, Verschmelzungen), Wortklassen, die überwiegend in der gesprochenen Sprache vorkommen (‚Diskursmarker‘), oder spezielle syntaktische Konstruktionen (z. B. *weil/obwohl/wobei/während* mit Verbzweitstellung oder Referenz-Aussage-Strukturen).¹ Behandelt werden also auffällige Einzelphänomene. Sie werden häufig als ‚ungrammatisch‘, ‚umgangssprachlich‘ oder ‚salopp‘ gekennzeichnet (vgl. Hennig 2001).

Die weiteren Einheiten der gesprochenen Sprache – funktionale Einheit, Gesprächsbeitrag und Gespräch – werden in der Regel nicht als Gegenstand der Grammatik angesehen und entsprechend nicht behandelt.

Üblicherweise ist die Unterscheidung von gesprochener und geschriebener Sprache für Grammatiken nicht zentral, weil sie beanspruchen, ‚die‘ Sprache (bzw. ‚das‘ Sprachsystem) als solche(s) zu beschreiben. (Dabei ist es eine hochinteressante Frage, wie man das tun kann, ohne auf mündlich oder schriftlich realisierte Sprache zu rekurrieren.) Dieser Anspruch bedeutet jedoch in der Regel, dass Grammatiken sich unter der Hand auf die Beschreibung konzeptionell schriftlicher Sprache beschränken.

Die Gegenstandsbestimmung und aktuelle Ausgestaltung von Grammatikschreibung verdankt sich:

- (1) der *Schriftlastigkeit des Sprachbewusstseins*,
- (2) der historisch bedingten *Überbewertung* und *Überrepräsentation* bestimmter Einheiten und
- (3) der Auffassung, dass gesprochene und geschriebene Sprache *Varianten einer Sprache* sind, die sich nur unwesentlich unterscheiden und denen *ein gemeinsames Sprachsystem* zugrunde liegt.

Zu (1): Die Schriftlastigkeit des Sprachbewusstseins und die gravierenden Konsequenzen dieser Schriftlastigkeit sind ein notorisch unterschätzter Punkt.

¹ Zu Referenz-Aussage-Strukturen s. Fiehler (2000).

In Gesellschaften mit einer ausgebauten Schriftsprache ist das *gesellschaftliche Sprachbewusstsein* weitgehend schriftsprachlich geprägt („written language bias“; Linell 1982). Die Vorstellungen darüber, was Sprache ist, leiten sich primär aus dem Umgang mit und der Reflexion von geschriebener Sprache her.

Die Gründe, warum die geschriebene und nicht die gesprochene Sprache das Sprachbewusstsein prägt, sind vielfältig. Es seien nur drei genannt:

- Die bei der Textproduktion auftretenden Probleme und Schwierigkeiten richten das Bewusstsein stark auf die Strukturen und Eigenschaften der geschriebenen Sprache. Die Leichtigkeit und der automatische Charakter des Sprechens hingegen bewirken, dass gesprochene Sprache nicht in gleicher Weise ins Zentrum der Aufmerksamkeit und des Sprachbewusstseins rückt.
- Die ‚Anschaulichkeit‘ und die Dauerhaftigkeit von Texten – im Gegensatz zur Hörbarkeit und Flüchtigkeit der gesprochenen Sprache – sind die Grundlage für ihre objektmäßige Präsenz und haben seit jeher die Reflexion schriftlicher Texte systematisch begünstigt.
- Geschriebene Sprache wird gesellschaftlich als wichtiger erachtet und höher bewertet als gesprochene. Entsprechend groß ist der Aufwand, der für den Schriftspracherwerb und die Schulung der Schreibfähigkeiten getrieben wird. So besitzt in der schulischen Erziehung die Schriftsprache eindeutig das Primat. Schriftspracherwerb und das Erstellen schriftlicher Texte aller Art haben dort ein deutliches Übergewicht gegenüber der Schulung mündlicher Kommunikation.

Ferner gilt in vielen Kontexten nur das als verlässlich, bedeutsam und wertvoll, was geschrieben steht. Eine Folge der Bedeutsamkeit und Wertschätzung von geschriebener Sprache ist auch der Aufwand, der für die Aufbewahrung von Schriftprodukten (Bibliotheken, Archive) betrieben wird. Dem steht nichts Vergleichbares für die gesprochene Sprache gegenüber.

Die geschriebene Sprache prägt aber nicht nur das allgemeine Sprachbewusstsein, sondern in gleicher Weise auch die *wissenschaftliche Erforschung und Beschreibung* von Sprache. Das „written language bias“ betrifft dabei sowohl den Untersuchungsgegenstand als auch die Kategorien zur Analyse von Sprache.

Zentraler *Untersuchungsgegenstand* der wissenschaftlichen Beschäftigung mit Sprache waren in der Geschichte bisher (außer wenn es explizit um die Lautlichkeit der Sprache ging) de facto schriftliche Texte oder Beispielsätze, die auf der Grundlage eines schriftsprachlich geprägten Bewusstseins schriftnah produziert werden. Nur sie waren – als Texte – dauerhaft gegeben und so einer wiederholten Betrachtung und detaillierten Analyse zugänglich. Wissenschaftliche Sprachreflexion war also – allein schon aus Gründen der Ver-

füßbarkeit des Gegenstandes – weitgehend Reflexion von Texten und damit von Schriftlichkeit.

Die Schrift- und Textlastigkeit der Sprachwissenschaft betrifft aber nicht nur die Daten, die als Untersuchungsgegenstand dienen, sondern auch die *Analyse- und Beschreibungskategorien* als Untersuchungsinstrumente. Die überwiegende Zahl der linguistischen Kategorien wurde in der und für die Analyse geschriebener Texte entwickelt und dann in Grammatiken zu einem relativ festen Satz von Analyse- und Beschreibungskategorien kanonisiert. Beispiele für solche Kategorien sind ‚Satz‘, ‚Wort‘, ‚Anakoluth‘, ‚Elision‘ etc. Diese grammatischen Beschreibungskategorien sind – wie alle Kategorien – funktional ihrem Gegenstand angepasst, und das heißt der Analyse und Beschreibung von geschriebener Sprache.

Die zentralen Kategorien zur Analyse von Schriftlichkeit sind nicht abstrakt oder äußerlich, sondern sie manifestieren und materialisieren sich in der *Form der Schriftlichkeit*. Sie sind dort vergegenständlicht und jeder Blick auf einen Text führt sie vor Augen. So wird das ‚Wort‘ (was schriftsprachgeschichtlich keineswegs immer so war) durch die Spatien sichtbar, der ‚Satz‘ durch die Großschreibung am Anfang und den abschließenden Punkt, der ‚Nebensatz‘ durch das Komma, das ‚Substantiv‘ durch seine Großschreibung (zumindest in der deutschen Schriftsprache) etc. Diese Kategorien werden im Entwicklungsprozess der Schriftsprache als – sich verändernde – Form der Schriftlichkeit ausgearbeitet. Einmal entwickelt, ist die Aktivierung und Anwendung dieser Kategorien Voraussetzung jeder korrekten Textproduktion. U. a. dies ist es, was den Schriftspracherwerb und das Schreiben so schwierig macht. Dass diese Kategorien bei jedem Akt des Schreibens aktiviert und angewendet werden müssen, verdeutlicht noch einmal, welch intensiv prägenden Einfluss sie auf das Sprachbewusstsein haben müssen.

Die auf die Schriftsprache ausgerichteten Analyse- und Beschreibungskategorien sind darüber hinaus das einzige voll entwickelte Kategoriensystem. Ein Kategoriensystem, das in ähnlicher Weise funktional auf die gesprochene Sprache zugeschnitten wäre, existiert im Moment nur in Ansätzen.

So sind das schriftsprachlich dominierte Sprachbewusstsein und die für die Schriftsprache entwickelten Analyse- und Beschreibungskategorien zwangsläufig die Grundlage für das Verständnis und die Erkenntnis von gesprochener Sprache: Gesprochene Sprache wird durch die Brille der geschriebenen wahrgenommen, sie ist das Modell für das Verständnis von Mündlichkeit.

Das ‚Übersehen‘ der gesprochenen Sprache hat aber noch einen anderen Grund: Anders als die geschriebene Sprache ist die gesprochene ein flüchtiger Gegenstand, was seine Untersuchbarkeit lange Zeit einschränkt hat und seine Untersuchung in besonderer Weise schwierig gestaltet: Entweder ist man auf die Erinnerung an Äußerungen oder Gespräche angewiesen, oder

aber es bedarf technischer Möglichkeiten der Konservierung von Äußerungen und Gesprächen.

Die Erinnerung ist ein bekanntermaßen unzuverlässiges Mittel. Die Verarbeitung und Speicherung von Gesprochenem ist im normalen Kommunikationsprozess inhaltlich-thematisch ausgerichtet. Bestimmte Phänomene werden – ohne entsprechende Schulung – gar nicht wahrgenommen oder treten nicht in den Fokus der Aufmerksamkeit. Äußerungen und Gespräche können schon nach kurzer Zeit nicht mehr im Wortlaut wiedergegeben werden, sondern sind nur noch inhaltlich zusammengefasst präsent.

Die Entwicklung und Verbreitung entsprechender *technischer Geräte* zur Konservierung und Reproduktion von Gesprächen und Interaktionen, also von Tonbandgeräten, Kassettenrekordern und Videokameras, ist deshalb eine wesentliche Voraussetzung für eine detaillierte wissenschaftliche Untersuchung von mündlicher Kommunikation. Setzt man eine bestimmte Ausgereiftheit und Verbreitung solcher Geräte voraus, kann man sagen, dass sie erst seit den 60er Jahren des 20. Jahrhunderts gegeben ist.

Eine zweite wesentliche Voraussetzung für die Untersuchung gesprochener Sprache ist die Entwicklung von Verfahren zur *Verschriftlichung (Transkription)* konservierter Gespräche. Transkriptionen ermöglichen eine Vergegenwärtigung und ‚Betrachtung‘ der Äußerungen und Gespräche, wie sie allein durch das Abhören der Aufzeichnung nicht zu erreichen ist. Die Entwicklung solcher Transkriptionssysteme für sprachwissenschaftliche Zwecke erfolgte Hand in Hand mit dem Einsatz der genannten Geräte. Erst durch das Zusammenspiel von reproduzierbaren Aufnahmen auf der einen Seite und Transkriptionen auf der anderen wird gesprochene Sprache in einem hinreichenden Detaillierungsgrad untersuchbar und erst von diesem Zeitpunkt an kann sie überhaupt zu einem ernsthaften und gleichwertigen wissenschaftlichen Untersuchungsgegenstand werden.

Zu (2): Historisch hat sich eine Grammatikschreibung herausgebildet, bei der das geschriebene Wort und der geschriebene Satz im Mittelpunkt stehen. Nimmt man aber die oben genannten grundlegenden Einheiten ernst, so hat Grammatik zum einen deutlich mehr zu beschreiben, als sie gegenwärtig tut, und zum anderen werden auch größere Unterschiede zwischen einer Grammatik gesprochener und geschriebener Sprache sichtbar. So hat eine Grammatik der gesprochenen Sprache Einheiten zu beschreiben, die bisher in keiner Weise Gegenstand der Grammatikschreibung waren (Gespräch, Gesprächsbeitrag, funktionale Einheiten als Bestandteile des Gesprächsbeitrags).

Zu (3): Die Auffassung, dass gesprochene und geschriebene Sprache Varianten einer Sprache sind, die sich nur unwesentlich unterscheiden und denen ein gemeinsames Sprachsystem zugrunde liegt, verdankt sich der *langue*-Orientierung der Sprachwissenschaft. Diese Orientierung verdeckt den Blick

auf die funktionalen und distributionellen Unterschiede zwischen gesprochener und geschriebener Sprache und die aus ihnen resultierenden Unterschiede in den sprachlich-kommunikativen Mitteln. Zum einen handelt es sich dabei um qualitative Unterschiede, wenn bestimmte Phänomene nur mündlich oder nur schriftlich auftreten, zum anderen um quantitative Unterschiede, wenn es sich um unterschiedliche Häufigkeiten der Verwendung bzw. des Vorkommens handelt.

These 6: Besonderheiten einer Grammatik gesprochener Sprache

Für eine Grammatik gesprochener Sprache stellen sich eine Reihe von besonderen Aufgaben und Problemen: Diese betreffen (1) die Multimodalität mündlicher Verständigung, (2) den prozessualen Charakter von Gesprächen, (3) die Kategorienentwicklung, (4) die Gegenstandsbestimmung und (5) die Differenzbeschreibung von gesprochener Sprache.

Anmerkungen zu These 6:

Zu (1) *Multimodalität der Verständigung*: Will man geschriebene Sprache beschreiben, so kann man sich auf die verbale Dimension beschränken. Nicht so bei der gesprochenen Sprache. Mündliche Verständigung geschieht, wenn sie unter den Bedingungen wechselseitiger Wahrnehmung erfolgt, gleichzeitig und parallel auf verschiedenen Ebenen: Sie ist multimodal. Im Verständigungsprozess wirken die auf visuellen Wahrnehmungen und Schlüssen basierende *wahrnehmungs- und inferenzgestützte Kommunikation* mit der *körperlichen Kommunikation* und der *verbalen Kommunikation* in spezifischer Weise zusammen.

Will man mündliche Verständigung in ihrer Spezifik und besonderen Regelmäßigkeit beschreiben, darf man sich also nicht auf das Gesprochene beschränken, sondern muss die verschiedenen Verständigungsebenen und ihr Zusammenwirken, also eben die Multimodalität mündlicher Verständigung, erfassen.

Zu (2) *Prozessualität des Gesprächs*: Grundlage für die grammatische Beschreibung geschriebener Sprache sind fertig vorliegende Produkte (Sätze, Texte). Wenn Texte produziert sind, gewinnen sie ein dauerhaftes Eigenleben, für das der Prozess ihrer Hervorbringung nicht wesentlich ist. Gegenstand der Beschreibung gesprochener Sprache hingegen sind nicht nur solche vorliegenden Produkte, also Transkripte, sondern zugleich auch als unausblendbare Voraussetzung und konstitutiver Bestandteil der Prozess der Interaktion, der Prozess der Hervorbringung, in dem sie entstehen: Mündliche Verständigung ist ein kooperativer Prozess, der in der Zeit abläuft und für den diese Prozessualität und Zeitlichkeit konstitutiv ist. Gesprochene Sprache

und Gespräche sind Resultat einer gemeinsamen Hervorbringung, die sich nach und nach in der Zeit entwickelt.

Die Analyse gesprochener Sprache erfordert deshalb vorrangig eine *Prozessorientierung* anstelle einer *Produktororientierung*, wie sie für die Analyse geschriebener Sprache charakteristisch ist. Für die Analyse gesprochener Sprache gilt das methodologische Postulat, die Produktion von Äußerungen bzw. die Entwicklung eines Gesprächs in ihrer zeitlichen Abfolge nachzuvollziehen. Nur so wird deutlich, welche Funktion einzelne Elemente im Prozess mündlicher Verständigung haben, und erst auf dieser Grundlage können dann strukturbezogene Kategorien sinnvoll gebildet werden.

Zu (3) *Kategorienentwicklung*: Die Beschreibung von gesprochener Sprache und Gesprächen erfordert in Teilen ein eigenständiges Kategorieninventar. Dies insbesondere für die Einheiten, zu denen es keine Entsprechung im Bereich der geschriebenen Sprache gibt (Laut, funktionale Einheit, Gesprächsbeitrag, Gespräch). Aber auch in den anderen Bereichen ist eine einfache Übernahme von Kategorien nicht möglich, weil die für die Analyse von geschriebener Sprache entwickelten Kategorien vorrangig produkt- und strukturorientiert sind. Sie sind funktional ihrem Gegenstand angepasst und taugen nur begrenzt zur Beschreibung gesprochener Sprache (Beispiele: Satz, Ellipse, ‚Linksherausstellung‘ etc.). Hier ist – wenn möglich – eine prozessorientierte Reinterpretation der Kategorien oder die Entwicklung gegenstandsangemessener (funktionaler) Kategorien erforderlich.

Obwohl davon auszugehen ist, dass gesprochene Sprache kategorial eigenständig beschrieben werden muss, kann ihre Spezifik jedoch in bestimmten Teilbereichen nur erfasst werden, indem die Verhältnisse in der gesprochenen Sprache mit denen in der geschriebenen Sprache verglichen werden. Bei diesem Vergleich lässt es sich nicht vermeiden, dass auch Beschreibungskonzepte und Beschreibungskategorien der geschriebenen Sprache verwendet werden (z. B. Konzept und Kategorie des Satzes) und dass von diesem Standpunkt aus die Besonderheiten der gesprochenen Sprache dann als Abweichung bzw. Negation gefasst werden (z. B. dass in der gesprochenen Sprache vielfältige Formen von ‚nichtsatzförmigen‘ Äußerungen zu beobachten sind). Die dabei verwendeten schriftsprachlichen Konzepte oder Kategorien sind jedoch für die gesprochene Sprache weder zentral, noch sind sie speziell auf ihre Beschreibung zugeschnitten. Zum Teil ist ihre Verwendung der Tatsache geschuldet, dass entsprechende Kategorien für die gesprochene Sprache noch nicht zur Verfügung stehen.

Zu (4) *Gegenstandsbestimmung*: Über die Frage, was zur Mündlichkeit bzw. gesprochenen Sprache zu rechnen ist und was nicht, mithin über ihre Grenzen, besteht keineswegs Konsens. Der Gegenstand wird wissenschaftlich durchaus unterschiedlich zugeschnitten. Im Grundsatz konkurrieren

dabei zwei Sichtweisen: eine medial-extensionale und eine prototypisch-graduierende.

Die *medial-extensionale* Sichtweise versteht alle die Formen als Mündlichkeit, bei denen Verständigung in irgendeiner Weise mittels gesprochener Sprache erfolgt. Demnach gehört das Verlesen von Nachrichten oder das Besprechen eines Anrufbeantworters ebenso dazu wie die kommunikative Koordination von Tätigkeiten oder der Kaffeeklatsch. Die Medialität ist das alleinige Kriterium dafür, was zur Mündlichkeit gerechnet wird. Ein häufiges Interesse in diesem Zusammenhang ist, verschiedene Formen mündlicher Kommunikation zu differenzieren.

Beim *prototypisch-graduierenden* Zugang wird den verschiedenen Formen mündlicher Verständigung eine Gewichtung aufgeprägt: Bestimmte Formen sind deutlicher, besser, genuiner etc. mündlich als andere. In der Regel stehen dabei die interaktiv realisierten Formen im Mittelpunkt. Diese Sichtweise beinhaltet die Möglichkeit, 'schlechtere' Fälle, die – als wesentlich definierte – Merkmale nicht aufweisen, aus der Betrachtung oder – weitergehend – als keine 'echten' Fälle von Mündlichkeit aus dem Gegenstandsbereich auszuschließen.

Zu (5) *Differenzbeschreibung*: Gesprochene Sprache ist ein Oppositionsbegriff. Er setzt als Kontrast das Konzept der geschriebenen Sprache voraus. Die Beschreibung gesprochener Sprache impliziert so den Vergleich mit geschriebener. Viele Eigenschaften gesprochener Sprache lassen sich nur in ihrer Differenz zu den Verhältnissen im Bereich der geschriebenen Sprache erfassen. Das 'written language bias' führt dabei zu einem Denken und Vergleichen aus der Perspektive der geschriebenen Sprache. Die Verhältnisse in der geschriebenen Sprache werden als der Normalfall angesehen und demgegenüber Abweichungen in der gesprochenen Sprache konstatiert.

These 7: Gesprächsbeiträge und funktionale Einheiten

Das Konzept des Satzes ist für eine Grammatik gesprochener Sprache nicht von zentraler Bedeutung. An seine Stelle treten als grundlegende Einheiten der Gesprächsbeitrag und funktionale Einheiten, aus denen sich die Gesprächsbeiträge aufbauen.

Anmerkungen zu These 7:

Die Frage, was dem Satz der geschriebenen Sprache in der gesprochenen Sprache entspricht, oder allgemeiner: was auf dieser Ebene Grundeinheiten mündlicher Kommunikation sind, hat eine Reihe verschiedener Antworten gefunden.

Zwar wurde lange Zeit als Folge der Schriftlastigkeit des Sprachbewusstseins postuliert, dass die gesprochene Sprache in gleicher Weise wie die geschriebene aus Sätzen besteht, je intensiver jedoch die Auseinandersetzung mit gesprochensprachlichen Daten wurde, desto mehr tauchten Zweifel an dieser Annahme auf. Die Übertragung der Kategorie ‚Satz‘ auf die Analyse gesprochener Sprache wurde zunehmend kritisiert und mit der Aufforderung verbunden, nach angemesseneren Kategorien zu suchen.

Als Konsequenz dieser Kritik wurde eine Reihe alternativer Konzepte für grundlegende Einheiten gesprochener Sprache vorgeschlagen:

- Sprachliche/kommunikative Handlung; Sprechakt; kommunikative Minimalement (z. B. Austin 1962, Searle 1969, Zifonun et al. 1997)
- Äußerungseinheit (z. B. Rath 1976, 1985, 1990 und 1997)
- Äußerung (z. B. Rehbein 1995)
- Intonation unit (z. B. Chafe 1988)
- *turn* und *turn constructional unit* (z. B. Sacks/Schegloff/Jefferson 1974; Ford/Fox/Thompson 1996; Selting 1996, 1998)

Ohne dies im Einzelnen hier zeigen zu können,² muss festgestellt werden, dass die Debatte um Grundeinheiten der mündlichen Kommunikation – als Folge der ursprünglichen Orientierung am Satz – zu *strukturell-formorientiert* verlaufen ist und dass so die *funktionale Gliederung* der Beiträge, an der sich die Beteiligten primär orientieren, zu wenig beachtet worden ist.

Im Folgenden wird davon ausgegangen, dass der *Gesprächsbeitrag* (*turn*) eine wesentliche Einheit der mündlichen Kommunikation ist. Nimmt man ernst, dass sich mündliche Verständigung in Gesprächen vollzieht, so scheint es angemessen, als nächstniedrige Einheit unterhalb des Gesprächs das anzusetzen, was die Gesprächsbeteiligten in zeitlicher Abfolge hierzu beitragen. Dies wirft aber zugleich die Frage auf, aus was für Elementen der Beitrag seinerseits besteht.

Gesprächsbeiträge bestehen aus funktionalen Einheiten. *Funktionale Einheiten* sind die kleinsten Bestandteile des Beitrags, denen die Gesprächsbeteiligten im Prozess der Produktion und Rezeption der Beiträge eine *separate Funktion im und für den Kommunikationsprozess* zuschreiben können. Die Untereinheiten des Beitrags werden damit primär funktional, nicht syntaktisch oder prosodisch bestimmt. Die Identifizierung funktionaler Einheiten macht von syntaktischen und prosodischen Mitteln (syntaktischen Gestaltschlüssen, geschlossenen Intonationskonturen) unterstützenden Gebrauch, sie ist aber weder allein noch primär von ihnen abhängig. Ferner wird die Bestimmung funktionaler Einheiten an die Perspektive der Gesprächsbeteiligten

² Vgl. Fiehler et al. (2004: 175ff.).

gebunden: Sobald sie Teilen eines Beitrags solche Funktionen zuschreiben können, identifizieren sie diese Segmente als funktionale Einheiten.

Mit funktionalen Einheiten werden bestimmte *Aufgaben im Kommunikationsprozess* bearbeitet. So können z. B. die Aufgaben ‚jemanden grüßen‘, ‚einen Gesprächspartner adressieren‘, ‚eine Aussage machen‘, ‚eine Bewertung äußern‘, ‚antworten‘, ‚eine Redewiedergabe ankündigen‘, ‚einen Grund nennen‘ etc. durch funktionale Einheiten realisiert werden. Systematisiert man diese funktionalen Einheiten, so handelt es sich zum einen um bestimmte Sprechhandlungen, zum anderen um Aktivitäten, die den Kommunikationsprozess organisieren und strukturieren, und zum dritten um Aktivitäten, mit denen die Funktionen anderer funktionaler Einheiten expliziert werden.

Nicht alle funktionalen Einheiten haben den gleichen Status: Zu unterscheiden sind potenziell selbstständige funktionale Einheiten, assoziierte funktionale Einheiten und projizierende funktionale Einheiten.

Potenziell selbstständige funktionale Einheiten können für sich alleine einen Beitrag bilden. *Assoziierte funktionale Einheiten* sind nicht alleinstehend möglich. Sie erfordern eine Trägereinheit, von der sie abhängig sind; nur mit ihr zusammen bilden sie einen Beitrag. *Projizierende funktionale Einheiten* sind ebenfalls nicht selbstständig. Sie machen eine weitere funktionale Einheit erwartbar, die von der projizierenden Einheit syntaktisch abhängig ist. Nur beide zusammen stellen einen vollständigen Beitrag dar.

Besteht ein Beitrag aus mehreren funktionalen Einheiten, so ist er *mehrgliedrig*. Die funktionalen Einheiten realisieren jede für sich eine eigenständige kommunikative Funktion. Das Erkennen dieser Funktionen ermöglicht das Separieren dieser Einheiten im Gesprächsbeitrag. Die Mehrgliedrigkeit kann zum einen dadurch zustande kommen, dass mehrere potenziell selbstständige Einheiten – entsprechend den je konkreten kommunikativen Absichten – kombiniert werden. So besteht die folgende Äußerung aus drei potenziell selbstständigen funktionalen Einheiten: einer reaktiven bewertenden Stellungnahme, einer Aussage und einer Frage:

Gut. Das war gestern, aber was machen wir morgen?

Zum anderen können einer potenziell selbstständigen funktionalen Einheit andere, die von ihr abhängig sind, assoziiert werden. In der folgenden Äußerung sind einer potenziell selbstständigen Aussage eine Adressierung und eine Bedingung (in Form eines Nebensatzes bzw. einer Parenthese) assoziiert:

Hans, wir können das Ergebnis im nächsten Jahr wiederholen, wenn die Rahmenbedingungen gleich bleiben.

bzw.:

Hans – gleiche Rahmenbedingungen vorausgesetzt – können wir das Ergebnis im nächsten Jahr wiederholen.

Zum dritten kann die Mehrgliedrigkeit daraus resultieren, dass eine funktionale Einheit, die nicht selbstständig ist, projektive Kraft besitzt und eine andere erwartbar macht, die von ihr abhängt. In den folgenden Äußerungen projiziert ein Operator³ eine Aussage bzw. die Ankündigung einer Redewiedergabe die Redewiedergabe:

Kurz, wir haben den Gegner deutlich unterschätzt.

Er meinte, wir werden mit unserer Initiative keinen Erfolg haben.

Die Analyse von mehrgliedrigen Beiträgen gibt Aufschluss über typische Grundstrukturen („Baupläne“) von Beiträgen. Häufig sind z. B. dreigliedrige Beiträge, deren erste funktionale Einheit rückbezüglich auf den vorhergehenden Beitrag eingeht, deren zweite Einheit den kommunikativen Kern des Beitrags enthält und deren dritte Einheit eine hörersteuernde Aktivität darstellt:

Stimmt, aber die Leistung hätte durchaus noch besser sein können, nicht.

Sprachlichen Einheiten sehr unterschiedlichen Umfangs und Typs kann eine Funktion im und für den Kommunikationsprozess zugeschrieben werden. D. h. funktionale Einheiten können, was die verbalen und/oder nonverbalen Mittel bzw. Elemente angeht, sehr unterschiedlich gefüllt sein, sofern diese verschiedenen Füllungen nur geeignet sind, eine äquivalente Funktionszuschreibung zu ermöglichen. Eine Geste, ein einzelnes Wort, eine (Nominal-/Präpositional-)Phrase, ein freistehender „Nebensatz“ oder eine vollständige Aussage in Satzform können so funktional äquivalent sein und eine funktionale Einheit bilden.

Was an sprachlichen Mitteln bzw. Elementen notwendig ist, damit der Hörer eine Funktionszuschreibung vornehmen kann, ist dabei keineswegs konstant, sondern hängt von einer Reihe von Faktoren ab, vor allem davon, was die Kommunizierenden als gemeinsames Wissen, das keiner expliziten Versprachlichung bedarf, voraussetzen (können). So variiert es stark, wie viel und was an sprachlichen Mitteln in einer konkreten Situation erforderlich ist, um etwas für den Hörer als eine funktionale Einheit bestimmten Typs hinreichend deutlich machen zu können. Im Prozess der Identifizierung einer funktionalen Einheit spielen die sprachlichen Mittel und das jeweilige Wissen in spezifischer Weise zusammen: Je mehr an gemeinsamem Wissen vorhanden ist, desto weniger muss tendenziell sprachlich (verbal und nonverbal) expliziert werden, um funktionale Einheiten kommunikativ zu realisieren und erkennen zu können. Es gibt also eine – vom als gemeinsam unterstellten Wissen abhängige – Varianz der sprachlichen Explizitheit bei der kommunikativen Realisierung funktionaler Einheiten. Diese Interdependenz von ge-

³ Zu Operator-Skopos-Strukturen s. Barden/Elstermann/Fiehler (2001) und Fiehler et al. (2004).

meinsamem Wissen und sprachlichem Material ist der rationale Kern dessen, was als Elliptizität gesprochener Sprache diskutiert wird. Hierbei wird nicht wirklich etwas ausgelassen – dies erscheint nur so, wenn man die für schriftliche Kommunikation notwendigen Standards der sprachlichen Explizitheit voraussetzt und auf gesprochene Sprache überträgt –, sondern es ist aufgrund gemeinsamen Wissens und der situativen Präsenz von Sachverhalten lediglich nicht notwendig, bestimmte Elemente explizit zu versprachlichen. Dies ist auch der Fall, wenn im Vorgängerkontext realisierte lexikalische Einheiten, syntaktische Strukturen oder Äußerungen als präsent vorausgesetzt werden und auf ihnen in Form von (Konstruktions-)Übernahmen oder Fortsetzungen aufgebaut wird, ohne dass sie noch einmal explizit wiederholt werden.

Die Möglichkeit, eine Funktionszuschreibung vorzunehmen, ist also nicht davon abhängig, dass die sprachlichen Mittel Satzform haben. Es gibt keine in der Sache liegende Verbindung zwischen Satzform und funktionalen Einheiten. Gleichwohl besitzen viele funktionale Einheiten (genauer: das sprachliche Material, dem eine bestimmte kommunikative Funktion zugeschrieben wird) Satzform. Welche funktionalen Einheiten regelmäßig als Sätze realisiert werden und bei welchen dies nur im Ausnahmefall geschieht, bedarf der gesonderten Untersuchung.

Die Segmentierung von Beiträgen in funktionale Einheiten erfolgt von den Gesprächsbeteiligten online im Prozess der Rezeption der Beiträge. Dies hat eine Reihe von Konsequenzen. So können die Beteiligten – bedingt durch den nicht abgeschlossenen Stand der Realisierung des Beitrags – vorschnelle oder falsche Funktionszuschreibungen vornehmen, die auf der Grundlage des folgenden sprachlichen Materials revidiert werden müssen. Umgekehrt ist es möglich, schon vor der (vollständigen) Realisierung des für eine spezifische Handlung notwendigen sprachlichen Materials diese Handlung zu antizipieren und ggf. darauf zu reagieren.

Der Prozess der Identifizierung von Handlungen und der Segmentierung von Beiträgen in entsprechende funktionale Einheiten, wie Gesprächsbeteiligte ihn beim Verstehen vollziehen, kann an realisierten Beiträgen, die dokumentiert vorliegen, *analytisch (nach)vollzogen* werden, indem der Analysierende – in analoger Weise wie die Gesprächsbeteiligten – sich auf der Grundlage seines sprachlich-kommunikativen Wissens und seiner Interaktionskompetenz bemüht, den Handlungssinn der Beiträge zu verstehen und zu interpretieren. Die *Methodik* der Identifizierung elementarer Bestandteile besteht also darin, zunächst Beiträge abzugrenzen, in ihnen dann Handlungen zu identifizieren und den Beitrag entsprechend in funktionale Einheiten zu segmentieren. Dies gilt wohlgemerkt nur für den analytischen Prozess. Bei der online-Prozessierung, bei der auch funktionale Einheiten identifiziert

werden, steht – wie schon verdeutlicht – das Ende von Beiträgen und damit die Gesamtheit ihrer funktionalen Einheiten noch nicht fest.

Ein Kriterium für die Validität der Segmentierung von Beiträgen in funktionale Einheiten besteht darin, dass verschiedene Personen (vor die Aufgabe gestellt, dies zu tun) einen Beitrag als analytische Leistung übereinstimmend in Einheiten aufteilen. Auf diese Weise wird rekonstruiert, was Personen intuitiv in einem Beitrag als kleinste abgrenzbare Handlungen und damit als funktionale Einheiten empfinden. Die Abgrenzung funktionaler Einheiten ist so letztlich eine empirische Frage, die am Maßstab ‚intersubjektive Übereinstimmung‘ entschieden wird.

3 Ausblick

Inzwischen ist die 7. Auflage der Duden-Grammatik erschienen, und das Kapitel zur gesprochenen Sprache, dessen konzeptionelle Grundlagen hier vorgestellt wurden, nachlesbar (Fiehler 2005). Wenngleich das Kapitel selbst wie auch seine Einbettung in die Gesamtgrammatik sicher in vielfacher Hinsicht verbesserungsfähig sind, kann es m. E. nicht positiv genug bewertet werden, dass nach der IDS-Grammatik (Zifonun et al. 1997) mit ihrem Abschnitt C: „Zur Grammatik von Text und Diskurs“ nun ein zweiter Schritt getan ist, um die gesprochene Sprache systematisch in die Grammatikschreibung mit einzubeziehen.

Literatur

- Austin, John L. (1962): *How to do Things with Words*. Oxford: Clarendon Press.
- Barden, Birgit/Elstermann, Mechthild/Fiehler, Reinhard (2001): Operator-Skopus-Strukturen in gesprochener Sprache. In: Liedtke, Frank/Hundsnurscher, Franz (Hrsg.): *Pragmatische Syntax*. Tübingen: Niemeyer, 197–233.
- Behagel, Otto (1899): *Geschriebenes Deutsch und gesprochenes Deutsch*. Wiederabgedruckt in: Behagel, Otto (1967): *Von deutscher Sprache*. Wiesbaden: Sändig, 11–34.
- Chafe, Wallace L. (1988): Linking intonation units in spoken English. In: Haiman, John/Thompson, Sandra A. (Hrsg.): *Clause combining in grammar and discourse*. Amsterdam/Philadelphia: Benjamins, 1–27.
- Fiehler, Reinhard (2000): Über zwei Probleme bei der Untersuchung gesprochener Sprache. In: *Sprache und Literatur* 31, 23–42.

- Fiehler, Reinhard (2005): *Gesprochene Sprache*. In: Duden. Die Grammatik. 7. Auflage. Mannheim: Dudenverlag, 1175–1256.
- Fiehler, Reinhard/Barden, Birgit/Elstermann, Mechthild/Kraft, Barbara (2004): *Eigenschaften gesprochener Sprache*. Tübingen: Narr.
- Ford, Cecilia E./Fox, Barbara A./Thompson, Sandra A. (1996): Practices in the construction of turns: The „TCU“ revisited. In: *Pragmatics* 6:3, 427–454.
- Hennig, Mathilde (2001): *Welche Grammatik braucht der Mensch? Grammatikführer für Deutsch als Fremdsprache*. München: Iudicium.
- Linell, Per (1982): *The written language bias in linguistics*. Linköping: University of Linköping (Studies in Communication 2).
- Rath, Rainer (1976): *Gesprochenes Deutsch und seine Gliederung*. In: Beiträge zu den Fortbildungskursen des Goethe-Instituts für ausländische Deutschlehrer an Schulen und Hochschulen. München: Goethe-Institut, 63–76.
- Rath, Rainer (1985): *Geschriebene und gesprochene Form der heutigen Standardsprache*. In: Besch, Werner/Reichmann, Oskar/Sonderegger, Stefan (Hrsg.): *Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung*. 2. Halbbd. Berlin/New York: de Gruyter, 1651–1663 (Handbücher zur Sprach- und Kommunikationswissenschaft 2,2).
- Rath, Rainer (1990): „Satz“ und „Äußerungseinheit“. Syntaktische und interaktive Struktur in der Sprache? In: Leupold, Eynar/Petter, Yvonne (Hrsg.): *Interdisziplinäre Sprachforschung und Sprachlehre. Festschrift für Albert Raasch zum 60. Geburtstag*. Tübingen: Narr, 197–216.
- Rath, Rainer (1997): „Äußerungseinheit“ oder „möglicher Satz“? In: *Deutsche Sprache* 25, 1–20.
- Rehbein, Jochen (1995): *Segmentieren*. Unveröff. Manuskript (Verbmobil Memo 64. Okt. 1995).
- Sacks, Harvey/Schegloff, Emanuel A./Jefferson, Gail (1974): A simplest systematics for the organization of turn-taking for conversation. In: *Language* 50, 696–735.
- Searle, John R. (1969): *Speech Acts. An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- Selting, Margret (1996): On the interplay of syntax and prosody in the constitution of turn-constructive units and turns in conversation. In: *Pragmatics* 6:3, 357–388.
- Selting, Margret (1998): TCUs and TRPs: The Construction of Units in Conversational Talk. In: *Interaction and Linguistic Structure* No. 4.
- Zifonun, Gisela/Hoffmann, Ludger/Strecker, Bruno et al. (1997): *Grammatik der deutschen Sprache*. 3 Bände. Berlin/New York: de Gruyter (Schriften des Instituts für deutsche Sprache 7).